

Limitations of Sums of Bounded Read Formulas

Purnata Ghosal

Department of Computer Science and Engineering, IIT Madras, Chennai, India
purnata@cse.iitm.ac.in

B. V. Raghavendra Rao

Department of Computer Science and Engineering, IIT Madras, Chennai, India
bvrr@cse.iitm.ac.in

Abstract

Proving super polynomial size lower bounds for various classes of arithmetic circuits computing explicit polynomials is a very important and challenging task in algebraic complexity theory. We study representation of polynomials as sums of weaker models such as read once formulas (ROFs) and read once oblivious algebraic branching programs (ROABPs). We prove:

- (1) An exponential separation between sum of ROFs and read- k formulas for some constant k .
- (2) A sub-exponential separation between sum of ROABPs and syntactic multilinear ABPs.

Our results are based on analysis of the partial derivative matrix under different distributions. These results highlight richness of bounded read restrictions in arithmetic formulas and ABPs.

Finally, we consider a generalization of multilinear ROABPs known as strict-interval ABPs defined in [Ramya-Rao, MFCS2019]. We show that strict-interval ABPs are equivalent to ROABPs upto a polynomial size blow up. In contrast, we show that interval formulas are different from ROFs and also admit depth reduction which is not known in the case of strict-interval ABPs.

2012 ACM Subject Classification Theory of computation → Algebraic complexity theory

Keywords and phrases Algebraic Complexity Theory, Arithmetic Circuits, Lower Bounds

Digital Object Identifier 10.4230/LIPIcs.CVIT.2016.23

1 Introduction

Polynomials are one of the fundamental mathematical objects and have wide applications in Computer Science. Algebraic Complexity Theory aims at a classification of polynomials based on their computational complexity. In his seminal work, Valiant [39] laid foundations of Algebraic Complexity Theory and popularized arithmetic circuits as a natural model of computation for polynomials. He proposed the permanent polynomial perm_n :

$$\text{perm}_n = \sum_{\pi \in S_n} \prod_{i=1}^n x_{i\pi(i)},$$

as the primary representative of intractability in algebraic computation. In fact, Valiant[39] conjectured that the complexity of computing perm_n by arithmetic circuits is different from that of the determinant function which is now known as Valiant's hypothesis.

One of the important offshoots of Valiant's hypothesis is the arithmetic circuit lower bound problem: prove a super polynomial lower bound on the size of an arithmetic circuit computing an explicit polynomial of polynomial degree. Here, an explicit polynomial is one whose coefficients are efficiently computable. Baur and Strassen [5] obtained a super linear lower bound on the size of any arithmetic circuit computing the sum of powers of variables. This is the best known size lower bound for general classes of arithmetic circuits.

Lack of improvements in the size lower bounds for general arithmetic circuits lead the community to investigate restrictions on arithmetic circuits. Restrictions considered in the literature can be broadly classified into two categories: syntactic and semantic. Syntactic



© Purnata Ghosal and B. V. Raghavendra Rao;
licensed under Creative Commons License CC-BY

42nd Conference on Very Important Topics (CVIT 2016).

Editors: John Q. Open and Joan R. Access; Article No. 23; pp. 23:1–23:20

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

restrictions considered in the literature include restriction on fan-out i.e., arithmetic formulas, restriction on depth i.e., bounded depth circuits [14, 15, 35], and the related model of algebraic branching programs. Semantic restrictions include monotone arithmetic circuits [18, 41, 37], homogeneous circuits [9], multilinear circuits [30] and noncommutative computation [26].

Grigoriev and Razborov [15] obtained an exponential lower bound for the size of a depth three arithmetic circuit computing the determinant or and permanent over finite fields. In contrast, only almost cubic lower bound is known over infinite fields [20]. Explaining the lack of progress on proving lower bounds even in the case of depth four circuits, Agrawal and Vinay [1] showed that an exponential lower bound for the size of depth four circuits implies Valiant’s hypothesis over any field. This lead to intense research efforts in proving lower bounds for the size of constant depth circuits, the reader is referred to an excellent survey by Saptharishi et al. [34] for details.

Recall that a polynomial $p \in \mathbb{F}[x_1, \dots, x_n]$ is said to be *multilinear* if every monomial in p with non-zero coefficient is square-free. An arithmetic circuit is said to be multilinear if every gate in the circuit computes a multilinear polynomial. Multilinear circuits are natural models for computing multilinear polynomials. Raz [31] obtained super polynomial lower bounds on the size of multilinear formulas computing the determinant or permanent. Further, he gave a super polynomial separation between multilinear formulas and circuits [30]. In fact, Raz [31] considered a syntactic version of multilinear circuits known as *syntactic multilinear* circuits. An arithmetic circuit C is said to be syntactic multilinear, if for every product gate $g = g_1 \times g_2$ the sub-circuits rooted at g_1 and g_2 are variable disjoint. The syntactic version has an advantage that the restriction can be verified by examining the circuit whereas there is no efficient algorithm for testing if a circuit is multilinear or not. Following Raz’s work, there has been significant interest in proving lower bounds on the size of syntactic multilinear circuits. Exponential separation of constant depth multilinear circuits is known [11], while the best known lower bound for unbounded depth syntactic multilinear circuits is only almost quadratic [2].

An *Algebraic Branching Program* (ABP) is a model of computation for polynomials that generalize arithmetic formulas and were studied by Ben-Or and Cleve [6] who showed that ABPs of constant width are equivalent to arithmetic formulas. Nisan [26] proved exponential size lower bound for the size of an ABP computing the permanent when the variables are non-commutative. It is known that polynomial families computed by ABPs are the same as families of polynomials computed by skew circuits, a restriction of arithmetic circuits where every product gate can have at most one non-input gate as a predecessor [23]. Further, skew arithmetic circuits are known to characterize the complexity of determinant [38]. Despite their simplicity compared to arithmetic circuits, the best known lower bound for size of ABPs is only quadratic [21, 10]. Even with the restriction of syntactic multilinearity, the best known size lower bound for ABPs is only quadratic [16]. However, a super polynomial separation between syntactically multilinear formulas and ABPs is known [13].

Proving super quadratic size lower bounds for syntactic multilinear ABPs (smABPs for short) remains a challenging task. Given that there is no promising approach yet to prove super quadratic size lower bounds for smABPs, it is imperative to consider further structural restrictions on smABPs and formulas to develop finer insights into the difficulty of the problem. Following the works in [27, 29, 28], we study syntactic multilinear formulas and smABPs with restrictions on the number of reads of variables and the order in which variables appear in a smABP.

Models and Results: (1) **Sum of ROFs:** A *read-once formula* (ROF) is a formula where every variable occurs exactly once as a leaf label. ROFs are syntactic multilinear by definition and have received wide attention in the literature. Volkovich [40] gave a complete characterization of polynomials computed by ROFs. Further, Minahan and Volkovich [24] obtained a complete derandomization of the polynomial identity testing problem on ROFs. While most of the multilinear polynomials are not computable by ROFs [40], sum of ROFs, denoted by $\Sigma \cdot \text{ROF}$ is a natural model of computation for multilinear polynomials. Shpilka and Volkovich showed that a restricted form of $\Sigma \cdot \text{ROF}$ requires linear summands to compute the monomial $x_1 x_2 \cdots x_n$. Further, Mahajan and Tawari [22] obtained a tight lower bound on the size of $\Sigma \cdot \text{ROF}$ computing an elementary symmetric polynomial. Ramya and Rao [29] obtained an exponential lower bound on the number of ROFs required to compute a polynomial in VP. In this article, we improve the lower bound in [29] to obtain an exponential separation between read- k formulas and $\Sigma \cdot \text{ROF}$ for a sufficiently large constant k . Formally, we prove:

► **Theorem 1.** *There is constant $k > 0$ and a family of multilinear polynomials f_{PRY} computable by read- k formulas such that if $f_{\text{PRY}} = f_1 + f_2 + \cdots + f_s$, where f_1, \dots, f_s are ROFs, then $s = 2^{\Omega(n)}$.*

(2) **Sum of ROABPs:** A natural generalization of ROFs are read-once oblivious branching programs (ROABPs). In an ROABP, a layer reads at most one variable and every variable occurs in exactly one layer. Arguments in [26] imply that any ROABP computing the permanent and determinant requires exponential size. Kayal et al. [19] obtain an exponential separation between the size of ROABPs and depth three multilinear formulas. In [27], an exponential lower bound for the sum of ROABPs computing a polynomial in VP is given. We improve this bound to obtain a super polynomial separation between sum of ROABPs and smABPs:

► **Theorem 2.** *There is a multilinear polynomial family \hat{f} computable by smABPs of polynomial size such that if $\hat{f} = f_1 + \dots + f_s$, each $f_i \in \mathbb{F}[X]$ being computable by a ROABP of size $\text{poly}(n)$, then $s = \exp(\Omega(n^\epsilon))$ for some $\epsilon < 1/500$.*

(3) **Strict-interval ABPs and Interval formulas:** It may be noted that any sub-program of a ROABP computes a polynomial in an interval $\{x_i, \dots, x_j\}$ of variables for some $i < j$. A natural generalization of ROABPs would be to consider smABPs where every sub-program computes a polynomial in some interval of variables, while a variable can occur in multiple layers. These are known as *interval ABPs* and were studied by Arvind and Raja [4] who obtained a conditional lower bound for the size of interval ABPs. Ramya and Rao [28] obtained an exponential lower bound for a special case of interval ABPs known as *strict-interval ABPs*. We show that strict-interval ABPs are equivalent to ROABPs upto polynomial size:

► **Theorem 3.** *The class of strict-interval ABPs is equivalent to the class of ROABPs.*

Finally, we examine the restriction of intervals in syntactic multilinear formulas. We show that unlike ROFs, interval formulas can be depth reduced (Theorem 28).

Related Work: To the best of our knowledge, Theorem 11 is the first exponential separation between bounded read formulas and $\Sigma \cdot \text{ROF}$. Prior to this, only a linear separation between bounded read formulas and $\Sigma \cdot \text{ROF}$ was known [3].

Ramya and Rao [29] obtain an exponential separation between $\Sigma \cdot \text{ROF}$ and multilinear VP. Our result is an extension of this result for the case of a simpler polynomial computable by bounded read formulas. Mahajan and Tawari [22] obtain tight linear lower bound for $\Sigma \cdot \text{ROF}$ computing an elementary symmetric polynomial.

Kayal, Nair and Saha [19] obtain a separation between ROABPs and multilinear depth three circuits. The authors define a polynomial, efficiently computed by set multilinear depth three circuits, that has an exponential size ROABP computing it. This polynomial can be expressed as a sum of three ROFs. Later, Ramya and Rao [27] obtain a sub-exponential lower bound against the model of $\Sigma \cdot \text{ROABP}$ computing the polynomial defined by Raz and Yehudayoff [33]. Dvir et al. [13] obtain a super-polynomial lower bound on the size of syntactic multilinear formulas computing a polynomial that can be efficiently computed by smABPs. We use the polynomial defined by [13] and adapt their techniques to obtain a separation between smABPs and $\Sigma \cdot \text{ROABP}$.

Organization of the Paper: Section 2 contains basic definitions of the models of computations, concepts and explicit polynomials used in the rest of the paper. The rest of the sections each describe results with respect to a particular bounded-read model. Section 3 describes the lower bound on the $\Sigma \cdot \text{ROF}$ model and Section 4 describes the lower bound on the $\Sigma \cdot \text{ROABP}$ model which follows using the same arguments as in the work of Dvir et al. [13]. Section 5 shows that strict-interval ABP is a fresh way to look at ROABPs since the two models are equivalent. In Section 6 we see that Brent's depth reduction result ([8]) holds for the class of interval formulas.

2 Preliminaries

In this section, we present necessary definitions and notations. For more details, reader is referred to excellent surveys by Shpilka and Yehudayoff [36] and by Saptharishi et al. [34].

Arithmetic Circuits: Let $X = \{x_1, \dots, x_n\}$ be a set of variables. An arithmetic circuit C over a field \mathbb{F} with input X is a directed acyclic graph (DAG) where the nodes have in-degree zero or two. The nodes of in-degree zero are called input gates and are labeled by elements from $X \cup \mathbb{F}$. Non-input gates of C are called internal gates and are labeled from $\{+, \text{times}\}$. Nodes of out degree zero are called output gates. Typically, a circuit has a single output gate. Every gate v in C naturally computes a polynomial $f_v \in \mathbb{F}[X]$. The polynomial computed by C is the polynomial represented at its output gate. The *size* of a circuit denoted by $\text{size}(C)$, is the number of gates in it, and *depth* is the length of the longest root to leaf path in C , denoted by $\text{depth}(C)$. An *arithmetic formula* is a circuit where the underlying undirected graph is a tree. For a gate v in C , let $\text{var}(v)$ denote the set of all variables that appear as leaf labels in the sub-circuit rooted at v .

Multilinear polynomials are polynomials such that in every monomial, the degree of a variable is either 0 or 1. Multilinear circuits, where every gate in the circuit computes a multilinear polynomial, are natural models of computation for multilinear polynomials. A circuit C is said to be *syntactic multilinear* if for every product gate $v = v_1 \times v_2$ in C , we have $\text{var}(v_1) \cap \text{var}(v_2) = \emptyset$. By definition, a syntactic multilinear circuit is also multilinear and computes a multilinear polynomial.

An arithmetic formula F is said to be a *read-once formula* (ROF in short) if every input variable in X labels at most one input gate in F .

Algebraic branching program (ABP in short) is a model of computation of polynomials defined as analogous to the branching program model of computation for Boolean functions. An ABP P is a layered DAG with layers L_0, \dots, L_m such $L_0 = \{s\}$ and $L_m = \{t\}$ where s is the start node and t is the terminal node. Each edge is labeled by an element in $X \cup \mathbb{F}$. The output of the ABP P is the polynomial $p = \sum_{\rho \text{ is a } s \text{ to } t \text{ path}} \text{wt}(\rho)$, where $\text{wt}(\rho)$ is the

product of edge labels in the path ρ . Further, for any two nodes u and v let $[u, v]_P$ denote the polynomial computed by the subprogram P_{uv} of P with u as the start node and v as the terminal node. Let X_{uv} denote the set of variables that appear as edge labels in the subprogram P_{uv} . The size of an ABP P , denoted by $\text{size}(P)$ is the number of nodes in P .

In a *syntactic multilinear* ABP (smABP), every s to t path reads any input variables at most once. An ABP is *oblivious* if every layer reads at most one variable. A *read-once oblivious ABP* (ROABP) is an oblivious syntactic multilinear ABP where every variable can appear in at most one variable i.e., for every i , there is at most one layer j_i such that x_i occurs as a label on the edges from L_{j_i} to L_{j_i+1} .

An interval on the set $\{1, \dots, n\}$ with end-points $i, j \in [n]$, can be defined as $I = [i, j]$, $i < j$, where $I = \{\ell \mid i, j \in [n], i \leq \ell \leq j\}$. An *interval of variables* X_{ij} is defined such that $X_{ij} \subseteq \{x_\ell \mid \ell \in I, I = [i, j]\}$, where I is an interval on the set $\{1, \dots, n\}$. For an ordering $\pi \in S_n$, we define a π -interval of variables, $X_{ij} \subseteq \{x_{\pi(i)}, x_{\pi(i+1)}, \dots, x_{\pi(j)}\}$. In [4], Arvind and Raja defined a sub-class of syntactically multilinear ABPs known as *interval ABPs* and proved lower bounds against the same. Later, [28] defined a further restricted version of interval ABPs, denoted by strict-interval ABPs, defined as follows.

► **Definition 4.** ([28]) A *strict interval ABP* P is a *syntactically multilinear ABP* where we have the following:

1. For any pair of nodes u and v in P , the indices of variables occurring in the sub-program $[u, v]_P$ is contained in some π -interval I_{uv} called the associated interval of $[u, v]_P$; and
2. for any pairs of sub-programs of the form $[u, v]_P$, $[v, w]_P$, the associated π -intervals of variables are disjoint, i.e., $I_{uv} \cap I_{vw} = \emptyset$.

It may be noted that in a strict interval ABP, intervals associated with each sub-program need not be unique. We assume that the intervals associated are largest intervals with respect to set inclusion such that condition 2 in the definition above is satisfied.

The Partial Derivative Matrix: We need the notion of partial derivative matrices introduced by Raz [31] and Nisan [26] as primary measure of complexity for multilinear polynomials. The partial derivative matrix of a polynomial $f \in \mathbb{X}$ defined based on a partition $\varphi : X \rightarrow Y \cup Z$ of the X into two parts. We follow the definition in [31]:

► **Definition 5.** (Raz [31]) Let $\varphi : X \rightarrow Y \cup Z$ be a partition of the input variables in two parts. Let $\mathcal{M}_Y, \mathcal{M}_Z$ be the sets of all possible multilinear monomials in the variables in Y and Z respectively. Then we construct the partial derivative matrix $M_{f\varphi}$ for a multilinear polynomial f under the partition φ such that the rows of the matrix are indexed by monomials $m_i \in \mathcal{M}_Y$, the columns by monomials $s_j \in \mathcal{M}_Z$ and entry $M_{f\varphi}(i, j) = c_{ij}$, c_{ij} being the coefficient of the monomial $m_i \cdot s_j$ in f . We denote by $\text{rank}_\varphi(f)$ the rank of the matrix $M_{f\varphi}$.

We call φ an equi-partition when $|X| = n$, n even and $|Y| = |Z| = n/2$.

Raz [31] showed the following fundamental property of rank_φ :

► **Lemma 6.** Let g and h be multilinear polynomials in $\mathbb{F}[X]$. Then, $\forall \varphi : X \rightarrow Y \cup Z$, we have the following.

Sub-additivity: $\text{rank}_\varphi(g + h) \leq \text{rank}_\varphi(g) + \text{rank}_\varphi(h)$, and

Sub-multiplicativity: $\text{rank}_\varphi(gh) \leq \text{rank}_\varphi(g) \times \text{rank}_\varphi(h)$,

In both the cases, equality holds when $\text{var}(g) \cap \text{var}(h) = \emptyset$.

Two Explicit Polynomials: Polynomials that exhibit maximum rank of the partial derivative matrix under all or a large fraction of equi-partitions can be thought of as *high complexity* or *hard* polynomials. We need two such families found in the literature.

Raz and Yehudayoff [33] defined a multilinear in VP. To define this polynomial we denote an interval $\{a \mid i \leq a \leq j, a \in \mathbb{N}\}$, $i, j \in \mathbb{N}$ by $[i, j]$, and consider the sets of variables $X = \{x_1, \dots, x_{2n}\}$, $W = \{w_{i,\ell,j}\}_{i,\ell,j \in [2n]}$. We denote it as the Raz-Yehudayoff polynomial and define it as follows.

► **Definition 7 (Raz-Yehudayoff polynomial, [33]).** *Let us consider $f_{ij} \in \mathbb{F}[X, W]$ defined over the interval $[i, j]$. For $i \leq j$, the polynomial f_{ij} is defined inductively as follows. If $j - i = 0$, then $f_{ij} = 0$. For $|j - i| > 0$,*

$$f_{ij} = (1 + x_i x_j) f_{i+1, j-1} + \sum_{\ell \in [i+1, j-2]} w_{i,\ell,j} f_{i,\ell} f_{\ell+1,j},$$

where we assume without loss of generality, lengths of $[i, \ell]$, $[\ell + 1, j]$ are even and smaller than $[i, j]$. We define $f_{1,2n}$ as the Raz-Yehudayoff polynomial f_{RY} .

Note that, f_{RY} can be defined over any subset $X' \subseteq X$ such that $|X'|$ is even, by considering the induced ordering of variables in X' and considering intervals accordingly. We denote this polynomial as $f_{RY}(X')$ for $X' \subseteq X$. Raz and Yehudayoff showed:

► **Proposition 8.** ([33]) *Let $\mathbb{G} = \mathbb{F}(W)$ be the field of rational functions over the field \mathbb{F} and the set of variables W . Then for every equi-partition $\varphi : X \rightarrow Y \cup Z$, $\text{rank}_\varphi(f_{RY}) = 2^{n/2}$.*

Dvir et al. [13] defined a polynomial that is hard i.e., full rank with respect to a special class of partitions called *arc-partitions*. Suppose $X = \{x_0, \dots, x_{n-1}\}$ be identified with the set $V = \{0, \dots, n-1\}$. For $i, j \in V$, the set $[i, j] = \{i, (i+1) \bmod n, (i+2) \bmod n, \dots, j\}$ is called the *arc* from i to j . An arc pairing is a distribution on the set of all pairings (i.e., perfect matchings) on V obtained in $n/2$ steps as follows. Assuming a pairing (P_1, \dots, P_t) constructed in $t < n/2$ steps, where $P_1 = (0, 1)$, $[L_t, R_t]$ is the interval spanned by $\cup_{i \in [t]} P_i$ and the random pair P_{t+1} is constructed such that

$$P_{t+1} = \begin{cases} (L_t - 2, L_t - 1) \text{ with probability } 1/3, \\ (L_t - 1, R_t + 1) \text{ with probability } 1/3, \\ (R_t + 1, R_t + 2) \text{ with probability } 1/3, \end{cases}$$

and therefore, $[L_{t+1}, R_{t+1}] = [L_t, R_t] \cup P_{t+1}$.

Given a pairing $\mathcal{P} = \{P_1, \dots, P_{n/2}\}$ of V , there are exactly $2^{n/2}$ partitions of X , by assigning $\varphi(x_i) \in Y$ and $\varphi(x_j) \in Z$ or $\varphi(x_i) \in Z$ and $\varphi(x_j) \in Y$ independently for each pair $(i, j) \in \mathcal{P}$. An *arc partition* is a distribution on all partitions obtained by sampling an arc pairing as defined above and sampling a partition corresponding to the pairing uniformly at random. We denote this distribution on partitions by \mathcal{D} . For a pairing $\mathcal{P} = \{P_1, \dots, P_{n/2}\}$ let $M_{\mathcal{P}}$ be the degree $n/2$ polynomial $\prod_{i=1}^{n/2} (x_{\ell_i} + x_{r_i})$ where $P_i = (\ell_i, r_i)$. Dvir et al. [13] defined the arc full rank polynomial $\hat{f} = \sum_{\mathcal{P} \in \mathcal{D}} \lambda_{\mathcal{P}} M_{\mathcal{P}}$, where $\lambda_{\mathcal{P}}$ is a formal variable. Dvir et al. [13] showed:

► **Proposition 9.** [13] *The polynomial \hat{f} can be computed by a polynomial size smABP and for every $\varphi \in \mathcal{D}$, $\text{rank}_\varphi(\hat{f}) = 2^{n/2}$ over a suitable field extension \mathbb{G} of \mathbb{F} .*

Now that we are familiar with most of the definitions required for an understanding of the results in this paper, we proceed to discuss our results.

3 Sum of ROFs

In [29], Ramya and Rao show an exponential lower bound for the sum of ROFs computing a polynomial in VP. While this establishes a super polynomial separation between $\Sigma \cdot \text{ROF}$ and syntactic multilinear formulas, it is interesting to see if this separation is exponential. In this section we obtain such an exponential separation. In fact, we show that there is an exponential separation between syntactic multilinear read- k formula and $\Sigma \cdot \text{ROF}$. We begin with the construction of a hard polynomial computable by a read- k formula for a large enough constant k .

A full rank Polynomial: Let $X = \{x_1, \dots, x_n\}$ be the set of input variables of the hard polynomial such that $4 \mid n$. Let $f_{\text{RY}}(X')$ to denote the Raz-Yehudayoff polynomial defined on the variable set X' of even size, where X' is an arbitrary subset of X .

Let $r = \Theta(1)$ be a sufficiently large integer factor of n such that r and n/r are both even. For $1 \leq i \leq n/r$, let $B_i = \{x_{(i-1)r+1}, \dots, x_{ir}\}$ and \mathcal{B} denote the partition $B_1 \cup B_2 \cup \dots \cup B_{n/r}$ of X . The polynomial f_{PRY} is defined as follows:

$$f_{\text{PRY}} = f_{\text{RY}}(B_1) \cdot f_{\text{RY}}(B_2) \cdots f_{\text{RY}}(B_{n/r}). \quad (1)$$

By definition of the polynomial f_{PRY} , it can be computed by a constant-width ROABP of polynomial size as well as by a read- k formula where $k = 2^{O(r)}$.

In order to prove a lower bound against a class of circuits computing the polynomial f_{PRY} , we consider the complexity measure of the rank of partial derivative matrix. Like in [30] and many follow-up results, we analyse the rank of the partial derivative matrix of f_{PRY} under a random partition. The reader might have already noticed that there are equi-partitions under which the $\text{rank}_\varphi(f_{\text{PRY}}) = 1$. Thus, we need a different distribution on the equi-partitions under which f_{PRY} has full rank with probability 1. In fact, under any partition φ , which induces an equi-partition on each of the variable blocks B_i , we have $\text{rank}_\varphi(f_{\text{PRY}}) = 2^{n/2}$, i.e., full rank. We define $\mathcal{D}_{\mathcal{B}}$ as the uniform distribution on all such partitions. Formally, we have:

► **Definition 10.** (Distribution $\mathcal{D}_{\mathcal{B}}$) The distribution $\mathcal{D}_{\mathcal{B}}$ is the distribution on the set of all equi-partitions $\hat{\varphi}$ of X obtained by independently sampling an equi-partition φ_i of each variable blocks B_i , for all i such that $1 \leq i \leq n/r$. We express $\hat{\varphi}$ as $\hat{\varphi} = \varphi_1 \circ \dots \circ \varphi_{n/r}$.

For any partition φ in the support of $\mathcal{D}_{\mathcal{B}}$, we argue that the polynomial f_{PRY} has full rank:

► **Observation 1.** For any $\varphi \sim \mathcal{D}_{\mathcal{B}}$, $\text{rank}_\varphi(f_{\text{PRY}}) = 2^{n/2}$ with probability 1.

Proof. Let us fix an equi-partition function $\hat{\varphi} \sim \mathcal{D}_{\mathcal{B}}$, $\hat{\varphi} : X \rightarrow Y \cup Z$. Let $t = r$. Considering $f_{\text{RY}}(X')$ where $|X'| = t$ and t is even, we can prove the partial derivative matrix of $f_{\text{RY}}(X')$ has rank $2^{n/2}$ under $\hat{\varphi}$ by induction on t . By definition of f_{RY} , for $t = 2$ we have $f_{\text{RY}} = 0$.

So, for the higher values of t , we see the term $(1 + x_1 x_t)$ and $f_{2,t-1}$ are variable disjoint, where $(1 + x_1 x_t)$ has rank ≤ 2 , and by the induction hypothesis, $f_{2,t-1}$ has rank $2^{t/2-1}$. Also, by induction hypothesis, for any ℓ , the ranks of partial derivative matrices of $f_{1,\ell}$ and $f_{\ell+1,t}$ are $2^{\ell/2}$ and $2^{(t-\ell)/2}$ respectively.

When $\hat{\varphi}(x_1) \in Y$ and $\hat{\varphi}(x_t) \in Z$, we set $w_{1,\ell,t} = 0$ for all $\ell \in [2, t-1]$ and $\text{rank}_{\hat{\varphi}}(f_{1,t}) = \text{rank}_{\hat{\varphi}}(1 + x_1 x_t) \cdot \text{rank}_{\hat{\varphi}} f_{2,t-1} = 2 \cdot 2^{(t/2-1)} = 2^{t/2}$. When $\hat{\varphi}(x_1) \in Y$ and $\hat{\varphi}(x_t) \in Y$, for an arbitrary $\ell \in [t]$ we set $w_{1,\ell,t} = 1$ and we have $\text{rank}_{\hat{\varphi}}(f_{1,t}) = \text{rank}_{\hat{\varphi}}(f_{1,\ell}) \cdot \text{rank}_{\hat{\varphi}}(f_{\ell+1,t}) = 2^{\ell/2}$, since $\hat{\varphi}$ is an equi-partition.

By sub-additivity of rank, and since B_i , $i \in [n/r]$ are disjoint sets of variables, we have $\text{rank}_{\hat{\varphi}}(f_{\text{PRY}}) = \prod_{i \in [n/r]} \text{rank}_{\hat{\varphi}}(f_{\text{RY}}(B_i)) = \prod_{i \in [n/r]} 2^{t/2} = 2^{tn/2r} = 2^{n/2}$. ◀

3.1 Rank Upper Bound on ROFs

In the following, we argue that the polynomial h cannot be computed by sum of ROFs of sub-exponential size. More formally,

► **Theorem 11.** *Let f_1, \dots, f_s be read-once polynomials such that $f_{\text{PRY}} = f_1 + f_2 + \dots + f_s$, then $s = 2^{\Omega(n)}$.*

We use the method of obtaining an upper bound on the rank of partial derivative matrix for ROFs with respect to a random partition developed by [29]. Though the argument in [29] works for an equi-partition sampled uniformly at random, we show their structural analysis of ROFs can be extended to the case of our distribution $\mathcal{D}_{\mathcal{B}}$. We begin with the notations used in [29] for the categorisation of the gates in a read-once formula F . (In this categorisation, the authors have only considered gates with at least one input being a variable.)

- Type- A: These are sum gates in F with both inputs variables in X .
- Type- B: Product gates in F with both inputs variables in X .
- Type- C: Sum gates in F where only one input is a variable in X .
- Type- D: Product gates in F where only one input is a variable in X .

Thus, type-D gates compute polynomials of the form $h \cdot x_i$ where $x_i \in X, h \in \mathbb{F}[X \setminus \{x_i\}]$ are the inputs to the type-D gate. Let a, b, c, d be the number of gates of type-A, B, C and D respectively. Let a'' be the number of Type A gates that compute a polynomial of rank 2 under an equi-partition φ , and a' be the number of Type-A gates that compute a polynomial of rank 1 under φ such that $a = a' + a''$.

The following lemma is an adaptation, for our distribution $\mathcal{D}_{\mathcal{B}}$, of the same lemma for the distribution of all equi-partitions on n variables from [29].

► **Lemma 12.** *Let $f \in \mathbb{F}[X]$ be an ROP, and φ be an equi-partition function sampled uniformly at random from the distribution $\mathcal{D}_{\mathcal{B}}$. Then with probability at least $1 - 2^{-\Omega(n)}$, $\text{rank}_{\varphi}(M_f) \leq 2^{n/2 - \Omega(n)}$.*

Proof. We first argue a rank upper bound for an arbitrary f_i . Let Φ_i be the formula computing f_i with gates of the types described as above. Let $\hat{\varphi} = \varphi_1 \circ \dots \circ \varphi_{n/r}$ sampled from the distribution \mathcal{D}_B uniformly at random.

We use the Lemma 3.1 from [29] which concludes that type-D gates do not contribute to the rank of a ROF.

► **Lemma 13.** [29, Lemma 3.1] *Let F be a ROF computing a read-once polynomial f and $\varphi : X \rightarrow Y \cup Z$ be an partition function on n variables. Then, $\text{rank}_{\varphi}(f) \leq 2^{a'' + \frac{2a'}{3} + \frac{2b}{3} + \frac{9c}{20}}$.*

Intuitively, Lemma 13 can be applied to a ROF F under a distribution $\hat{\varphi} \sim \mathcal{D}_B$ as follows. If there are a large number of type D gates (say αn , for some $0 \leq \alpha < 1$), then for any such equi-partition $\hat{\varphi}$, $\text{rank}_{\hat{\varphi}}(f) \leq 2^{(1-\alpha)n/2}$. A type C gate, too, contributes a small value (at most 2) to the rank compared to gates of types A and B. Thus, without loss of generality, we assume that the number of type C and D gates is at most αn . Now our analysis proceeds as in [29], only differing in the estimation of a'' , a' under an equi-partition $\hat{\varphi} \sim \mathcal{D}_B$.

Let (P_1, \dots, P_t) be a pairing induced by the gates of types A and B (i.e., the two inputs to a gate of type A or B form a pair). There can be at most $n/2$ pairs, but since we have αn gates of type C and D for some $0 \leq \alpha < 1$, we assume $(1 - \alpha)n$ remaining type A and B

gates. Therefore, for $t = (n - \alpha n)/2$, $t \leq n/2$, we have the pairs P_1, \dots, P_t induced by the type A and B gates in Φ_i .

Now, considering the division of X into $B_1, \dots, B_{n/r}$, we can divide the pairs into two sets depending on whether a pair lies entirely within a block B_i , $i \in [n/r]$ or the pair has its members in two different blocks B_i and B_j for $i, j \in [n/r]$, $i \neq j$. We define these two sets as $W = \{P_i \mid P_i = (x, y), \exists \ell, x, y \in B_\ell\}$ for pairs lying within blocks and $A = \{P_i \mid P_i = (x, y), \exists j, k, j \neq k, x \in B_j, y \in B_k\}$ for pairs lying across blocks, where x, y are two arbitrary variables in X .

Each pair P_i can be monochromatic or bichromatic under the randomly sampled equipartition $\hat{\varphi}$ with the probability $\frac{1}{2}$. Presence of monochromatic edges will give us a reduction in the rank of f_i under $\hat{\varphi}$. The analysis on W and A is done separately as follows.

Analysing W , $|W| > t/2$:

Let $B_{i_1}, \dots, B_{i_\ell}$ be the blocks containing at least one pair from W , $\ell \leq n/r$. We want to estimate ℓ and count how many of these ℓ blocks have at least one monochromatic pair under $\hat{\varphi}$ from W .

For each B_i , $i \in [t]$, we define the Bernoulli random variable X_i such that,

$$X_i = \begin{cases} 1, & \text{if } \exists P \in W, P = (x, y), x, y \in B_i, \\ 0, & \text{otherwise.} \end{cases}$$

Let $\Pr[X_i = 1] = \Pr[\exists P \in W, P = (x, y), x, y \in B_i] = \epsilon$, for some $\epsilon > 0$.

Then we have $\mathbb{E}[X_i] = \epsilon$, and for $\mathcal{X} = X_1 + \dots + X_{n/r}$, $\mathbb{E}[\mathcal{X}] = \epsilon \cdot n/r$. By the Chernoff's bound defined in [25], we have,

$$\Pr[\mathcal{X} > 2\epsilon n/r] < \exp\left(\frac{-\epsilon n}{3r}\right).$$

Now we estimate ϵ as follows:

$$\begin{aligned} \epsilon &= \Pr[X_i = 1] = \Pr[\exists P \in W, P = (x, y), x, y \in B_i] \\ &= \Pr[x, y \in B_i \mid \exists P \in W, P = (x, y)] \\ &= \frac{\Pr[x, y \in B_i]}{\Pr[\exists P \in W, P = (x, y)]} \\ &\geq \Pr[x, y \in B_i] \text{ since } \Pr[\exists P \in W, P = (x, y)] \leq 1 \\ &= \frac{1}{r^2}. \end{aligned}$$

Therefore, $\Pr[\mathcal{X} > 2\epsilon n/r] < \exp\left(\frac{-\epsilon n}{3r}\right) \leq \exp(-\Omega(n))$, when r is a constant. This implies that at least $2/r^2$ fraction of the blocks have a pair entirely within them with probability $1 - \exp(-\Omega(n))$ and each of these pairs is monochromatic under $\hat{\varphi}$ with the constant probability $1/2$. This gives an upper bound on the rank of f_i ,

$$\text{rank}_{\hat{\varphi}}(f_i) \leq 2^{n/2 - n/r^3} = 2^{n/2 - \Omega(n)}.$$

Analysing A , $|A| > t/2$:

Since each pair of variables in A lies across two blocks, we create a graph $G = (V, E)$ where each $v_i \in V$ represents the block B_i and $E = \{(v_i, v_j) \mid (x, y) \in A, x \in B_i, y \in B_j, i \neq j\}$.

The graph G has maximum degree r since there can be at most r pairs with one member in a fixed block B_i . If the edges in E form a perfect matching M' in G , then under $\hat{\varphi}$, the

edges in E can be either bichromatic or monochromatic. We need to show there will be sufficient number of monochromatic edges to give a tight upper bound for $\text{rank}_{\hat{\varphi}}(f_i)$.

By a result in [7], any graph with maximum degree r has a maximal matching of size $m/(2r-1)$, where $|E| = m$. Since $|A| \geq t/2$, $m \geq t/2$ and hence the maximal matching is of size $t/2(2r-1) = \Omega(n)$ when r is a suitable constant. With probability $1/2$, an edge in the maximal matching is bichromatic. Hence, $\leq t/2$ number of the edges in the maximal matching are bichromatic with probability $1/2^{t/2} = O(\exp(n^{-1}))$. So, with the high probability of $1 - O(\exp(n^{-1}))$, more than half of the edges in the maximal matching are monochromatic, thus giving us the rank bound,

$$\text{rank}_{\hat{\varphi}}(f_i) \geq 2^{n/2-t/2} = 2^{n/2-\Omega(n)}.$$

◀

Given an upper bound on the rank of ROFs under a random partition from $\mathcal{D}_{\mathcal{B}}$, we now proceed to prove the Theorem 11 by showing a lower bound on the size of ROFs computing our hard polynomial h .

Proof. (Proof of Theorem 11) By Observation 1, the upper bound on the rank of ROFs given by Lemma 12 and the sub-additivity of rank, we have:

$$s \cdot 2^{n/2-\Omega(n)} \leq 2^{n/2} \implies s = 2^{\Omega(n)}.$$

◀

With this result, the relationship between the classes of polynomials computable by polynomial size ROFs, ROABPs and depth-3 multilinear circuits is clear. Since the class of smABPs of polynomial size is strictly smaller than the class of polynomial size multilinear circuits (as in the non-multilinear setting), in the next section we obtain a lower bound on the sum of ROABPs computing the explicit polynomial in [13], which is efficiently computable by smABPs.

4 A separation between Sum of ROABPs and smABPs

In this section we prove a sub-exponential lower bound against the size of sum of read-once oblivious ABPs computing the hard polynomial constructed in [13]. This shows a sub-exponential separation between syntactically multilinear ABPs and sum of ROABPs.

We prove the following theorem in this section:

► **Theorem 2.** *There is a multilinear polynomial family \hat{f} computable by smABPs of polynomial size such that if $\hat{f} = f_1 + \dots + f_s$, each $f_i \in \mathbb{F}[X]$ being computable by a ROABP of size $\text{poly}(n)$, then $s = \exp(\Omega(n^\epsilon))$ for some $\epsilon < 1/500$.*

Our aim is to give an upper bound on the maximum rank of ROABPs under an arc partition. We refer to the rank of the coefficient matrix of the sum of ROABPs against an arc-partition as the *arc-rank*. We analyze the arc-rank of the sum of ROABPs against an arc-partition to give a lower bound on the size of the sum necessary to compute \hat{f} .

Let us assume that n is even. In order to prove the lower bound, we need to estimate an upper bound on the arc-rank computed by a ROABP. We define the notion of F -arc-partition, F being a ROABP, as follows:

► **Definition 14.** Let us consider an arc partition Q constructed from a ROABP F in the following manner: Let the order of variables appearing in the ROABP be $x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}$, where $\sigma \in S_n$ is a permutation on n indices. Then, $Q = \{(x_{\sigma(i)}, x_{\sigma(i+1)}) \mid i \in [n], i \text{ is odd}\}$ is a F -arc-partition.

We assume $2K \mid n$. Let S_1, \dots, S_K be a K -coloring of the variable set X , where x_1, \dots, x_n are ordered according to the ROABP and for every $i \in [k]$, S_i contains the variables $x_{(i-1)n/K+1}, \dots, x_{in/K}$ according to that ordering. Then S_1, \dots, S_K is a K -partitioning of the pairs in the F -arc-partition Q . So pairs in Q are monochromatic, whereas the pairs $(P_1, \dots, P_{n/2})$ on which a random arc-partition Π sampled from \mathcal{D} is based, might cross between two colors.

Our analysis for the ROABP arc-rank upper bound follows along the lines of the analysis for the arc-rank upper bound given by [13] for syntactic multilinear formulas. For this analysis we define the set of violating pairs for each color c , $V_c(\Pi)$, that is defined as:

$$V_c(\Pi) = \{\Pi_t \mid |\Pi_t \cap S_c| = 1, t \in [n/2]\},$$

where $\Pi_1, \dots, \Pi_{n/2}$ are pairs in Π . The quantity $G(\Pi) = |\{c \mid |V_c(\Pi)| \geq n^{\frac{1}{1000}}\}|$, representing the number of colors with many violations, is similarly defined. We use the following lemma directly from [13]:

► **Lemma 15.** Let $K \leq n^{\frac{1}{100}}$, Π be the sampled arc-partition, and $G(\Pi)$ be as defined above. Then, we have, $\Pr_{\Pi \in \mathcal{D}}[G(\Pi) \leq K/1000] \leq n^{-\Omega(K)}$.

The following measure is used to compute the arc-rank upper bound for ROABPs.

► **Definition 16.** (Similarity function) Let φ be a distribution on functions $\mathcal{S} \times \mathcal{S} \rightarrow \mathbb{N}$, such that \mathcal{S} is the support of the distribution on arc-partitions, \mathcal{D} . Let P, Q be arc-partitions sampled independently and uniformly at random from \mathcal{D} . Then, $\varphi(Q, P) : \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{N}$ is the total number of common pairs between two arc-partitions Q and P .

We assume Q to be the F -arc-partition for the ROABP F . For a pair that is not common between Π and Q , we show both the variables in the pair is in the same partition, Y or Z with high probability.

► **Theorem 17.** Under an arc-partition Π sampled from \mathcal{D} uniformly at random, if $p \in \mathbb{F}[X]$ is the polynomial computed by a ROABP P , then, for the similarity function φ and $\delta > 0$,

$$\Pr_{\Pi \sim \mathcal{D}}[\varphi(\Pi, Q) \geq n/2 - n^\delta] \leq 2^{-o(n)}.$$

Proof Outline: Our argument is the same as [13]. It is being included here for completeness for the parameters here being somewhat different than [13].

In order to analyse the number of common pairs counted by φ , we consider the K -coloring of F and show that under a random arc-partition Π , the number of crossing pairs are large in number using Lemma 15. Then, we show, this results in large number of pairs having both elements in Y . In order to identify the colors with the high number of crossing pairs, a graphical representation of the color sets is used.

Proof. [13] construct the graph $H(\Pi)$, where each vertex is a color c such that $|V_c(\Pi)| \geq n^{\frac{1}{1000}}$, and vertices c and d have an edge connecting them if and only if $|V_c(\Pi) \cap V_d(\Pi)| \geq n^{\frac{1}{1500}}$. We know for any two colors $c, d \in [K]$, $|V_c(\Pi) \cap V_d(\Pi)| \leq n^{\frac{1}{1000}}$. So, by definition of $H(\Pi)$, the least degree of a vertex in $H(\Pi)$ is 1. Using this, [13] prove the following claim:

23:12 Limitations of Sums of Bounded Read Formulas

▷ **Claim 18.** Let the size of the vertex set of $H(\Pi)$, $V(H(\Pi))$, be M . For any subset U of $V(H(\Pi))$ size $N \geq M/2 - 1$, there is some color $h_{j+1}, j \in [N - 1]$ such that in the graph induced on all vertices except $\{h_1, \dots, h_j\}$, the degree of h_{j+1} is at least 1.

By Claim 18, we have $U \subseteq V(H(\Pi))$, $U = \{c_1, \dots, c_{M/2-1}\}$ such that this is the set of colours having high number of crossing pairs common with colors not in U . Considering the colors sequentially, given Π , we first examine the pairs crossing from color c_1 to other colors, then c_2 and so on. Therefore, to examine the event E_i for color c_i , we have to estimate $\Pr_{\Pi \sim \mathcal{D}}[E_i \mid E_1, \dots, E_{i-1}, \Pi]$.

Here, E_i is the event $|Y_{c_i} - |S_{c_i}|/2| \leq n^{\frac{1}{5000}}$, equivalently expressed as $|S_{c_i}|/2 - n^{\frac{1}{5000}} \leq Y_{c_i} \leq |S_{c_i}|/2 + n^{\frac{1}{5000}}$. But for an upper bound, it suffices to analyse the $n^{\frac{1}{1500}}$ crossing pairs from S_{c_i} to S_{c_j} instead of considering the entire set. Let the subset of Y_{c_i} constituted by one end of crossing pairs going to color c_j be P_{ij} . Each element x in a crossing pair $P_t = (x, w)$ is a binomial random variable in a universe of size $\geq n^{\frac{1}{1500} = s}$ with probability $1/2$ of being allotted to the subset Y of the universe. This event is independent of how the c_i colored element of other crossing pairs $P_{t'}$ are allotted. So, $|B_{ij}| = b_j$ is a hypergeometric random variable where B_{ij} contains all such $x \in Y$. By the properties of a hypergeometric distribution, $\Pr_{b_j}[b_j = a] = O(s^{-\frac{1}{2}}) = O(n^{\frac{-1}{3000}})$, where a is a specific value taken by the size of B_{ij} .

Applying the union bound over all colors c_j for the crossing pairs, and taking $b = \sum_{j \in U \setminus \{i\}} b_j$, we have:

$$\Pr_b[s/2 - n^{\frac{1}{5000}} \leq b \leq |S_{c_i}|/2 - n^{\frac{1}{5000}}] \leq 2n^{\frac{1}{5000}} O(n^{\frac{-1}{3000}}) = n^{-\Omega(1)}.$$

Therefore, $\Pr_{\Pi \sim \mathcal{D}}[E_i \mid E_1, \dots, E_{i-1}, \Pi] = n^{-\Omega(\delta)}$.

We want an upper bound for $\Pr[|Y_c - |S_c||/2| \leq n^{\frac{1}{5000}} \forall c \in [K]]$. We have calculated an upper bound for the colors in $[K]$ that were highly connected to each other in $H(\Pi)$. So, we can now estimate the total probability as follows:

$$\begin{aligned} & \Pr[|Y_c - |S_c||/2| \leq n^{\frac{1}{5000}} \forall c \in [K]] \\ &= \mathbb{E}[n^{-\Omega(G(P))} \mid G(P) > K/1000] + \mathbb{E}[n^{-\Omega(G(P))} \mid G(P) \leq K/1000] \\ &= \mathbb{E}[n^{-\Omega(G(P))} \mid G(P) > K/1000] + n^{-\Omega(K)} \text{ by Lemma 15} \\ &\leq n^{-\Omega(K)}. \end{aligned}$$

If we consider $\delta = 1/5000$, then:

$$\Pr_{\Pi \sim \mathcal{D}}[\varphi(\Pi, Q) \geq n/2 - n^\delta] \leq \Pr[|Y_c - |S_c||/2| \leq n^{\frac{1}{5000}} \forall c \in [K]] \leq n^{-\Omega(K)}$$

Now, in Lemma 15, $K \leq n^{\frac{1}{1000}}$.

$$\text{Hence, } \Pr_{\Pi \sim \mathcal{D}}[\text{rank}_\varphi(M(p_\Pi)) \geq 2^{n/2 - n^\delta}] \leq 2^{-cn^{\frac{1}{1000}} \log n} = 2^{-o(n)}. \quad \blacktriangleleft$$

Now, using the above Theorem 17, we can prove the lower bound on the size of the sum of ROABP, s .

Proof. (of Theorem 2) Since the polynomial f is such that each multiplicand is of the form $\lambda_e(x_u + x_v)$, if x_u, x_v are both mapped to the same partition Y or Z , it will reduce the rank of the partial derivative matrix by half. Hence, we have the following:

$$\Pr_{\Pi \sim \mathcal{D}}[\text{rank}_\varphi(M(f_\Pi)) \geq 2^{n/2 - n^\delta}] = \Pr_{\Pi \sim \mathcal{D}}[\varphi(\Pi, Q) \geq n/2 - n^\delta],$$

for some suitable $\delta > 0$.

$$\begin{aligned}
\Pr[\text{rank}(M(f_\Pi)) = 2^{n/2}] &\leq \Pr[\exists i \in [s], \text{rank}(M((f_i)_\Pi)) \geq 2^{n/2}/s] \\
&\leq \sum_{i=1}^s \Pr[\text{rank}(M((f_i)_\Pi)) \geq 2^{n/2}/s] \\
&\leq \sum_{i=1}^s \Pr[\text{rank}(M((f_i)_\Pi)) \geq 2^{n/2-n^\delta}] \text{ for some } \delta > 0 \\
&\leq s \cdot n^{-\Omega(n^{\frac{1}{1000}})} \\
\implies s &= 2^{\Omega(n^{\frac{1}{1000}} \log n)} = 2^{\Omega(n^{\frac{1}{500}})}.
\end{aligned}$$

◀

The difference in computational power of ROABPs and smABPs highlights the power of reads of variables. From their definition, strict-interval ABPs generalise ROABPs by reading an interval of variables in every sub-program instead of reading a subset of variables in a fixed order. However, in the following section, we note that reading in intervals do not lend more computational power, and that ROABPs and Strict-interval ABPs in fact compute the same class of polynomials.

5 Strict-Interval ABPs

A strict-interval ABP, defined in [28] (See Definition 4), is a restriction of the notion of interval ABPs introduced by [4]. In the original definition given by [28], every sub-program in a strict-interval ABP P is defined on a π -interval of variables for some order π , however, without loss of generality, we assume π to be the identity permutation on n variables. Therefore, an interval of variables $[i, j]$, $i < j$ here is the set $\{x_i, \dots, x_j\}$. In this section we show that strict-interval ABPs are equivalent to ROABPs upto a polynomial blow-up in size.

► **Theorem 19.** *The class of strict-interval ABPs is equivalent to the class of ROABPs.*

The proof of Theorem 19 involves a crucial observation that in a strict-interval ABP, variables are read in at most two orders and the nodes that correspond to paths that read in different orders can be isolated. We start with some observations on intervals in $[1, n]$ and the intervals involved in a strict interval ABP.

Let P be a strict-interval ABP over the variables $X = \{x_1, \dots, x_n\}$. For any two nodes u and v in P , let $I_{u,v}$ be the interval of variables associated with the sub-program of P with u as the start node and v as the terminal node. For two intervals $I = [a, b]$, $J = [c, d]$ in $[1, n]$, we say $I \preceq J$, if $b \leq c$. Note that any two intervals I and J in $[1, n]$ are comparable under \preceq if and only if either they are disjoint or the largest element in one of the intervals is the smallest element in the other. This defines a natural transitive relation on the set of all intervals in $[1, n]$. The following is a useful property of \preceq :

► **Observation 2.** *Let I, J and J' be intervals over $[1, n]$ such that $I \preceq J$ and $J' \subseteq J$. Then $I \preceq J'$.*

Proof. Let $I = [a, b]$, $J = [c, d]$ and $J' = [c', d']$. As $I \preceq J$, we have $b \leq c$. Further, since $J' \subseteq J$, we have $c \leq c'$ and $d' \leq d$. Therefore, $b \leq c'$ and hence $I \preceq J'$. ◀

We begin with an observation on the structure of intervals of the sub-programs of P . Let v be a node in P . We say v is an *ascending* node, if $I_{s,v} \preceq I_{v,t}$ and a *descending* node if $I_{v,t} \preceq I_{s,v}$.

► **Observation 3.** *Let P be a strict-interval ABP and v any node in P . Then, v is either ascending or descending and not both.*

Proof. Let $I = I_{s,v}$ and $J = I_{v,t}$. Since P is a strict-interval ABP, the intervals I and J are disjoint and hence either $I \preceq J$ or $J \preceq I$ as required. ◀

Consider any s to t path ρ in P . We say that ρ is *ascending* if every node in ρ except s and t is ascending. Similarly, ρ is called *descending* if every node in ρ except s and t is descending.

► **Lemma 20.** *Let P be a strict interval ABP and let ρ any s to t path in P . Then either ρ is ascending or descending.*

Proof. We prove that no s to t path in P can have both ascending and descending nodes. For the sake of contradiction, suppose that ρ has both ascending and descending nodes. There are two cases. In the first, there is an edge (u, v) in ρ such that u is an ascending node and v is a descending node. Let $I = I_{s,u}$, $J = I_{u,t}$, $I' = I_{s,v}$ and $J' = I_{v,t}$. Since $P_{s,u}$ is a sub-program of $P_{s,v}$, we have $I \subseteq I'$, similarly $J' \subseteq J$. By the assumption, we have $I \preceq J$ and $J' \preceq I'$. By Observation 2, we have $I \preceq J'$ and $J' \preceq I'$. By transitivity, we have $I \preceq I'$. However, by the definition of \preceq , I and I' are incomparable, which is a contradiction. The second possibility is u being a descending node and v being an ascending node. In this case, $J \preceq I$ and $I' \preceq J'$. Then, by Observation 2, we have $J' \preceq I$ as $J' \subseteq J$. Therefore, $J \preceq J'$ by the transitivity of \preceq , a contradiction. This completes the proof. ◀

Lemma 20 implies that the set of all non-terminal nodes of P can be partitioned into two sets such that there is no edge from one set to the other. Formally:

► **Lemma 21.** *Let P be an interval ABP. There exist two strict-interval ABPs P_1 and P_2 such that*

1. *All non-terminal nodes of P_1 are ascending nodes and all non-terminal nodes of P_2 are descending nodes; and*
2. *$P = P_1 + P_2$.*

Proof. Let P_1 be the sub-program of P obtained by removing all descending nodes from P and P_2 be the sub-program of P obtained by removing all ascending nodes in P . By Lemma 20, the non-terminal nodes in P_1 and P_2 are disjoint and every s to t path ρ in P is either a s to t path in P_1 or a s to t path in P_2 but not both. Thus $P = P_1 + P_2$. ◀

Next we show that any strict-interval ABP consisting only of ascending or only of descending nodes can in fact be converted into an ROABP.

► **Lemma 22.** *Let P be a strict-interval ABP consisting only of ascending nodes or only of descending nodes. Then the polynomial computed by P can also be computed by a ROABP P' of size polynomial in $\text{size}(P)$. The order of variables in P' is x_1, \dots, x_n if P has only ascending nodes and x_n, \dots, x_1 if P has only descending nodes.*

Proof. We consider the case when all non-terminal nodes of P are ascending nodes. Let ρ be any s to t path in P . We claim that the edge labels in ρ are according to the order x_1, \dots, x_n . Suppose that there are edges (u, v) and (u', v') occurring in that order in ρ such that (u, v) is labelled by x_i and (u', v') is labelled by x_j with $j < i$. Let $I' = I_{s,u'}$ and $J' = I_{u',t}$. Since $i \in I'$, $j \in J'$ and $I' \cap J' = \emptyset$, it must be the case that $J' \preceq I'$ and hence u' must be a descending node, a contradiction. This establishes that P is an one ordered ABP. By the equivalence between one ordered ABPs and ROABPs ([16], [17]), we conclude that

the polynomial computed by P can also be computed by a ROABP of size polynomial in the size of P .

The argument is similar when all non-terminal nodes of P are descending. In this case, we have $i < j$ in the above argument and hence $I' \preceq J'$, making u' an ascending node leading to a contradiction. This concludes the proof. \blacktriangleleft

A permutation π of $[1, n]$ naturally induces the order $x_{\pi(1)}, \dots, x_{\pi(n)}$. The *reverse* of π is the order $x_{\pi(n)}, x_{\pi(n-1)}, \dots, x_{\pi(1)}$. Since branching programs are layered, any multilinear polynomial computed by a ROABP where variables occur in the order given by π can also be computed by a ROABP where variables occur in the reverse of π .

► **Observation 4.** *Let P be a ROABP where variables occur in the order induced by a permutation π . The polynomial computed by P can also be computed by a ROABP of same size as P that reads variables in the reverse order corresponding to π .*

Proof. Let P' be the ROABP obtained by reversing the edges of P and swapping the start and terminal nodes. Since P is a layered DAG, there is a bijection between the set of all s to t paths in P and the set of all s to t paths in P' , where the order of occurrence of nodes and hence the edge labels are reversed. This completes the proof. \blacktriangleleft

The above observations immediately establish Theorem 19.

Proof of Theorem 19. Let P be a strict-interval ABP of size S computing a multilinear polynomial f . By Lemma 21 there are strict interval ABPs P_1 and P_2 such that P_1 has only ascending non-terminal nodes and P_2 has only descending non-terminal nodes such that $f = f_1 + f_2$ where f_i is the polynomial computed by P_i , $i \in \{1, 2\}$. By Lemma 22 and Observation 4, f_1 and f_2 can be computed by a ROABPs that read the variables in the order x_1, \dots, x_n . Then $f_1 + f_2$ can also be computed by an ROABP. It remains to bound the size of the resulting ROABP. Note that $\text{size}(P_i) \leq S$. A ROABP for f_i can be obtained by staggering the reads of P_i which blows up the size of the ABP by a factor of n ([16], [17]). Therefore size of the resulting ROABP is at most $2nS \leq O(S^2)$. \blacktriangleleft

Using Theorem 19, we can design the following white-box PIT for strict-interval ABPs.

► **Corollary 23.** *Given a strict-interval ABP P of size s , we can check whether the polynomial computed by P is identically zero in time $O(\text{poly}(S))$.*

Proof. The proof follows from Theorem 19 and the polynomial time white-box PIT algorithm given by [32] for non-commutative ABPs, since the variables in X are read only once, in a fixed order, in a ROABP. \blacktriangleleft

The notion of intervals of variables corresponding to every sub-program can be applied to formulas in the form of Interval Formulas, where every sub-formula corresponds to an interval. In the following section we explore how such a model can be used to generalize the model of ROFs, and in what ways it differs from ROFs.

6 Interval Formulas

We saw that strict-interval ABPs have the same computational power as ROABPs despite being seemingly a non-trivial generalization. It is naturally tempting to guess that a similar generalization of ROFs might yield a similar result. However, we observe that such a generalization of ROFs yields a class different from ROFs.

We introduce interval formulas as a generalization of read-once formulas. An interval on variable indices, $[i, j]$, $i < j$, is an interval corresponding to the set of variables $X_{ij} \subseteq X = \{x_1, \dots, x_n\}$, where $X_{ij} = \{x_p \mid x_p \in X, i \leq p \leq j\}$. Polynomials are said to be defined on the interval $[i, j]$ when the input variables are from the set X_{ij} . When there is no ambiguity, we refer to X_{ij} as an interval of variables $[i, j]$. Gates in a read-once formula F can also be viewed as reading an interval of variables according to an order π on the variables i.e., there is a permutation $\pi \in S_n$ such that every gate v in F is a sub-formula computing a polynomial on a π -interval of variables. Thus, interval formulas are a different generalization of read-once formulas where every gate v in the formula F reads an interval of variables in a fixed order.

We formally define interval formulas as follows:

► **Definition 24.** (*Interval Formulas*) An arithmetic formula F is an interval formula if for every gate g in F , there is an interval $[i, j]$, $i < j$ such that g computes a polynomial in X_{ij} and for every product gate $g = h_1 \times h_2$, the intervals corresponding to h_1 and h_2 must be non-overlapping.

Thus, if a product gate g in F defined on an interval $I = [i, j]$ takes inputs from gates g_1, \dots, g_t , then the gates g_1, \dots, g_t compute polynomials on disjoint intervals $[i, j_1], [j_1 + 1, j_2], \dots, [j_{t-1} + 1, j]$ respectively, where $\forall p, j_p < j_{p+1}$ and $i \leq j_p \leq j$. If g_1, g_2 , defined on intervals I_1, I_2 are input gates to a sum gate g' , then the interval I associated with g' is $I = I_1 \cup I_2$.

A quick observation is that interval formulas are different from ROFs:

► **Proposition 25.** The set of all polynomials computable by interval formulas is different from that of ROFs

Proof. By [40], the polynomial $x_1x_2 + x_2x_3 + x_1x_3$ is not an ROF. However, the expression $x_1x_2 + x_2x_3 + x_1x_3$ is itself an interval formula. ◀

In fact, interval formulas are universal, since any sum of monomials can be represented by an interval formula.

Our next observation is that the polynomial f_{PRY} defined in Section 3 can be computed by an interval formula.

► **Proposition 26.** The polynomial family f_{PRY} is computable by an interval formula of polynomial size.

Proof. Recall that $f_{\text{PRY}}(X) = f_{\text{RY}}(B_1) \cdot f_{\text{RY}}(B_2) \cdots f_{\text{RY}}(B_{n/r})$. Since each of the $f_{\text{RY}}(B_i)$ is a constant variate polynomial and the sum of product representation of any multilinear polynomial is an interval formula by definition, we have that $f_{\text{RY}}(B_i)$ is computable by an interval formula of constant size. This $f_{\text{PRY}}(X)$ has a polynomial size interval formula. ◀

It is not known if every ROF can be converted to a ROF of logarithmic depth. However, we argue, in the following that interval formulas can be depth-reduced efficiently.

We have the following depth reduction result for general arithmetic formulas given by [8], who showed that depth of any arithmetic formula can be reduced by allowing its size to be increased by a polynomial factor.

► **Theorem 27.** [8] Any polynomial p computed by an arithmetic formula of size s and depth d , can also be computed by a formula of size $\text{poly}(s)$ and depth $O(\log s)$.

We know that this reduction preserves multilinearity. However, we don't know if Theorem 27 can be modified to preserve the read- k property. We show that the depth reduction algorithm given by Theorem 27 preserves the interval property.

► **Theorem 28.** *Let $f \in \mathbb{F}[X]$ be a polynomial computed by an interval formula F of size s and depth d . Then f can also be computed by an interval formula of size $\text{poly}(s)$ and depth $O(\log s)$.*

Proof. We know that the underlying structure of any arithmetic formula is a tree. The proof by Brent crucially uses the fact that by the *tree-separator lemma* [12], we are guaranteed that there exists a tree-separator node g such that the sub-tree Φ of a formula Φ' of total size s , rooted at the node g , has size $\leq 2s/3$.

The construction by [8] proceeds as follows. We replace the gate g by a new formal variable y . Let the resulting polynomial computed by F be $f'(x_1, \dots, x_n, y)$, where $f(x_1, \dots, x_n) = f'(x_1, \dots, x_n, g)$ under the new substitution $y = g$. As f' is linear in y , we have

$$f'(x_1, \dots, x_n, y) = yf_1(x_1, \dots, x_n) + f_0(x_1, \dots, x_n),$$

where $f_0 = f'|_{y=0}$ and $f_1 = f'|_{y=1} - f'|_{y=0}$. Thus, f_0, f_1 can be computed by multilinear formulas of size less than $\text{size}(F)$. Now, recursively obtaining small-depth formulas for f_1, f_0 , we obtain a $O(\log s)$ depth formula computing f .

However, the above construction does not necessarily preserve the interval property, since the intervals of variables on which f_0, f_1 and g are defined, can be overlapping. We overcome this problem by expressing f_0, f_1 as products of polynomials over disjoint intervals, each of the intervals being disjoint to the interval corresponding to g .

We assume, without loss of generality, that the interval formula F corresponds to the interval $[1, n]$. Let the interval corresponding to g be $I_g = [i, j]$, $i < j$. Now, by definition of f_1 and f_0 , they are defined on the same interval of variables. We consider the intervals I_0, I_1 such that $I_0 \cup I_1 = [1, n] \setminus [i, j]$, $I_0 = [j+1, n]$ and $I_1 = [1, i-1]$. We express both f_0, f_1 as products of two polynomials on I_0 and I_1 respectively. As f_1 and g are multiplicatively related in F , we show that $f_1 = f_{1,1} \times f_{1,0}$ where $f_{1,1}$ is a polynomial on the interval I_1 and $f_{1,0}$ is a polynomial on the interval I_0 .

We consider the root to leaf (g) path ρ in the original formula F containing the node g . All the paths meeting ρ at a sum gate represent polynomials additively related to y i.e., contributing towards the computation of f_0 and not f_1 . For f_1 , we will analyze only the paths meeting ρ at product gates. Let us consider a product gate on ρ computing $h_1 \times h_2$, such that h_2 lies on ρ . Since I is contained in the interval corresponding to h_2 , the interval corresponding to h_1, I_{h_1} must be either fully contained in I_1 or I_0 .

Constructing an interval formula for f_1 : We ignore all sum gates on ρ computing $p_1 + p_2$, with p_2 on ρ , by substituting p_1 to zero. The resulting formula is F' . In any product gate computing $h_1 \times h_2$, where h_2 is on ρ , if $I_{h_1} \subset I_0$, we substitute h_1 by 1. We also substitute g by 1. The remaining formula F'_1 computes the polynomial $f^{(1)}$.

We repeat this process above, but this time, we substitute h_1 by 1 only when $I_{h_1} \subset I_1$. This remaining formula F'_2 computes $f^{(2)}$. By definition of f_1 , $f_1 = f^{(1)} \cdot f^{(2)}$. The interval corresponding to F'_1 is contained in I_1 , the interval corresponding to F'_2 is contained in I_0 .

Constructing an interval formula for f_0 : We ignore all product gates on ρ computing $h_1 \times h_2$, with h_2 on ρ , by substituting h_1 by 1. The resulting formula is \hat{F} .

In any sum gate computing $p_1 + p_2$, where p_2 is on ρ , if $I_{p_1} \subset I_0$, we substitute p_1 by 0. We also substitute g by 0. The remaining formula \hat{F}_1 computes the polynomial $p^{(1)}$.

We repeat this process from the beginning, but substitute p_1 by 0 only when $I_{p_1} \subset I_1$. This remaining formula \hat{F}_2 computes $p^{(2)}$. By definition of f_0 , $f_0 = p^{(1)} + p^{(2)}$. The interval corresponding to \hat{F}_1 is contained in I_1 , the interval corresponding to \hat{F}_2 is contained in I_0 .

Hence, we obtain $f = f^{(1)}f^{(2)}g + p^{(1)} + p^{(2)}$. The recursive relation for calculating depth is as follows: $\text{depth}(F) = \text{depth}(g) + 2 \implies \text{depth}(s) = \text{depth}(2s/3) + 2$, which yields a total depth of $O(\log s)$ for F . ◀

References

- 1 Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 67–75. IEEE Computer Society, 2008. doi:10.1109/FOCS.2008.32.
- 2 Noga Alon, Mrinal Kumar, and Ben Lee Volk. Unbalancing sets and an almost quadratic lower bound for syntactically multilinear arithmetic circuits. *Comb.*, 40(2):149–178, 2020. doi:10.1007/s00493-019-4009-0.
- 3 Matthew Anderson, Dieter van Melkebeek, and Ilya Volkovich. Derandomizing polynomial identity testing for multilinear constant-read formulae. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity, CCC 2011, San Jose, California, USA, June 8-10, 2011*, pages 273–282. IEEE Computer Society, 2011. doi:10.1109/CCC.2011.18.
- 4 Vikraman Arvind and S. Raja. Some lower bound results for set-multilinear arithmetic computations. *Chicago J. Theor. Comput. Sci.*, 2016, 2016. URL: <http://cjtcs.cs.uchicago.edu/articles/2016/6/contents.html>.
- 5 Walter Baur and Volker Strassen. The complexity of partial derivatives. *Theor. Comput. Sci.*, 22:317–330, 1983. doi:10.1016/0304-3975(83)90110-X.
- 6 Michael Ben-Or and Richard Cleve. Computing algebraic formulas using a constant number of registers. *SIAM J. Comput.*, 21(1):54–58, 1992. doi:10.1137/0221006.
- 7 Therese C. Biedl, Erik D. Demaine, Christian A. Duncan, Rudolf Fleischer, and Stephen G. Kobourov. Tight bounds on maximal and maximum matchings. *Discret. Math.*, 285(1-3):7–15, 2004. doi:10.1016/j.disc.2004.05.003.
- 8 Richard P. Brent. The parallel evaluation of general arithmetic expressions. *J. ACM*, 21(2):201–206, 1974. doi:10.1145/321812.321815.
- 9 Peter Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*, volume 7 of *Algorithms and computation in mathematics*. Springer, 2000.
- 10 Prerona Chatterjee, Mrinal Kumar, Adrian She, and Ben Lee Volk. A quadratic lower bound for algebraic branching programs. In Shubhangi Saraf, editor, *35th Computational Complexity Conference, CCC 2020, July 28-31, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 169 of *LIPIcs*, pages 2:1–2:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.CCC.2020.2.
- 11 Suryajith Chillara, Christian Engels, Nutan Limaye, and Srikanth Srinivasan. A near-optimal depth-hierarchy theorem for small-depth multilinear circuits. In Mikkel Thorup, editor, *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 934–945. IEEE Computer Society, 2018. doi:10.1109/FOCS.2018.00092.
- 12 Fan RK Chung. *Separator theorems and their applications*. Forschungsinst. für Diskrete Mathematik, 1989. URL: <http://www.math.ucsd.edu/~fan/mypaps/fanpap/117separatorthms.pdf>.
- 13 Zeev Dvir, Guillaume Malod, Sylvain Perifel, and Amir Yehudayoff. Separating multilinear branching programs and formulas. In Howard J. Karloff and Toniann Pitassi, editors, *Proceed-*

- ings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012, pages 615–624. ACM, 2012. doi:10.1145/2213977.2214034.
- 14 Dima Grigoriev and Marek Karpinski. An exponential lower bound for depth 3 arithmetic circuits. In Jeffrey Scott Vitter, editor, *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*, pages 577–582. ACM, 1998. doi:10.1145/276698.276872.
 - 15 Dima Grigoriev and Alexander A. Razborov. Exponential lower bounds for depth 3 arithmetic circuits in algebras of functions over finite fields. *Appl. Algebra Eng. Commun. Comput.*, 10(6):465–487, 2000. doi:10.1007/s002009900021.
 - 16 Maurice J. Jansen. Lower bounds for syntactically multilinear algebraic branching programs. In Edward Ochmanski and Jerzy Tyszkiewicz, editors, *Mathematical Foundations of Computer Science 2008, 33rd International Symposium, MFCS 2008, Torun, Poland, August 25-29, 2008, Proceedings*, volume 5162 of *Lecture Notes in Computer Science*, pages 407–418. Springer, 2008. doi:10.1007/978-3-540-85238-4_33.
 - 17 Maurice J. Jansen, Youming Qiao, and Jayalal Sarma. Deterministic black-box identity testing \mathbb{F}_p -ordered algebraic branching programs. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2010, December 15-18, 2010, Chennai, India*, pages 296–307, 2010. doi:10.4230/LIPIcs.FSTTCS.2010.296.
 - 18 Mark Jerrum and Marc Snir. Some exact complexity results for straight-line computations over semirings. *J. ACM*, 29(3):874–897, 1982. doi:10.1145/322326.322341.
 - 19 Neeraj Kayal, Vineet Nair, and Chandan Saha. Separation between read-once oblivious algebraic branching programs (roabps) and multilinear depth three circuits. In Nicolas Ollinger and Heribert Vollmer, editors, *33rd Symposium on Theoretical Aspects of Computer Science, STACS 2016, February 17-20, 2016, Orléans, France*, volume 47 of *LIPIcs*, pages 46:1–46:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPIcs.STACS.2016.46.
 - 20 Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. An almost cubic lower bound for depth three arithmetic circuits. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:6, 2016. URL: <http://eccc.hpi-web.de/report/2016/006>.
 - 21 Mrinal Kumar. A quadratic lower bound for homogeneous algebraic branching programs. *Comput. Complex.*, 28(3):409–435, 2019. doi:10.1007/s00037-019-00186-3.
 - 22 Meena Mahajan and Anuj Tawari. Sums of read-once formulas: How many summands are necessary? *Theor. Comput. Sci.*, 708:34–45, 2018. doi:10.1016/j.tcs.2017.10.019.
 - 23 Guillaume Malod and Natacha Portier. Characterizing valiant’s algebraic complexity classes. *J. Complex.*, 24(1):16–38, 2008. doi:10.1016/j.jco.2006.09.006.
 - 24 Daniel Minahan and Ilya Volkovich. Complete derandomization of identity testing and reconstruction of read-once formulas. *TOCT*, 10(3):10:1–10:11, 2018. doi:10.1145/3196836.
 - 25 Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005. doi:10.1017/CB09780511813603.
 - 26 Noam Nisan. Lower bounds for non-commutative computation (extended abstract). In Cris Koutsougeras and Jeffrey Scott Vitter, editors, *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing, May 5-8, 1991, New Orleans, Louisiana, USA*, pages 410–418. ACM, 1991. doi:10.1145/103418.103462.
 - 27 C. Ramya and B. V. Raghavendra Rao. Lower bounds for special cases of syntactic multilinear abps. In Lusheng Wang and Daming Zhu, editors, *Computing and Combinatorics - 24th International Conference, COCOON 2018, Qing Dao, China, July 2-4, 2018, Proceedings*, volume 10976 of *Lecture Notes in Computer Science*, pages 701–712. Springer, 2018. doi:10.1007/978-3-319-94776-1_58.
 - 28 C. Ramya and B. V. Raghavendra Rao. Lower bounds for multilinear order-restricted abps. In Peter Rossmanith, Pinar Hegghernes, and Joost-Pieter Katoen, editors, *44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, August 26-*

- 30, 2019, Aachen, Germany, volume 138 of *LIPICs*, pages 52:1–52:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPICs.MFCS.2019.52.
- 29 C. Ramya and B. V. Raghavendra Rao. Lower bounds for sum and sum of products of read-once formulas. *TOCT*, 11(2):10:1–10:27, 2019. doi:10.1145/3313232.
 - 30 Ran Raz. Separation of multilinear circuit and formula size. *Theory of Computing*, 2(6):121–135, 2006. doi:10.4086/toc.2006.v002a006.
 - 31 Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *J. ACM*, 56(2):8:1–8:17, 2009. doi:10.1145/1502793.1502797.
 - 32 Ran Raz and Amir Shpilka. Deterministic polynomial identity testing in non-commutative models. *Comput. Complex.*, 14(1):1–19, 2005. doi:10.1007/s00037-005-0188-8.
 - 33 Ran Raz and Amir Yehudayoff. Balancing syntactically multilinear arithmetic circuits. *Comput. Complex.*, 17(4):515–535, 2008. doi:10.1007/s00037-008-0254-0.
 - 34 Ramprasad Satharishi, Suryajith Chillara, and Mrinal Kumar. A survey of lower bounds in arithmetic circuit complexity. Technical report, 2016. URL: <https://github.com/dasarpmar/lowerbounds-survey/releases>.
 - 35 Amir Shpilka and Avi Wigderson. Depth-3 arithmetic circuits over fields of characteristic zero. *Computational Complexity*, 10(1):1–27, 2001. URL: <http://dx.doi.org/10.1007/PL00001609>, doi:10.1007/PL00001609.
 - 36 Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends® in Theoretical Computer Science*, 5(3–4):207–388, 2010. URL: <http://dx.doi.org/10.1561/04000000039>, doi:10.1561/04000000039.
 - 37 Srikanth Srinivasan. Strongly exponential separation between monotone VP and monotone VNP. *CoRR*, abs/1903.01630, 2019. URL: <http://arxiv.org/abs/1903.01630>, arXiv:1903.01630.
 - 38 Seinosuke Toda. Classes of arithmetic circuits capturing the complexity of computing the determinant. *IEICE Transactions on Information and Systems*, 75(1):116–124, 1992.
 - 39 Leslie G. Valiant. The complexity of computing the permanent. *Theor. Comput. Sci.*, 8:189–201, 1979. doi:10.1016/0304-3975(79)90044-6.
 - 40 Ilya Volkovich. Characterizing arithmetic read-once formulae. *TOCT*, 8(1):2:1–2:19, 2016. doi:10.1145/2858783.
 - 41 Amir Yehudayoff. Separating monotone VP and VNP. In Moses Charikar and Edith Cohen, editors, *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 425–429. ACM, 2019. doi:10.1145/3313276.3316311.